

DSA Academy, Recruitment Tech Task

Vinted Technical Task for the Entry Level position (IC1) to the Academy Programme

Technical Task at Vinted

Welcome to this Vinted Technical task! The goal of this assignment is to evaluate your skills required for an entry level position in DSA. Before we start, let's discuss some ground rules for this assignment:

- You should spend **4 hours** on this task
- The output should be your **analyses** (Python/R/SQL) and a single **pdf file**:
 - Please include in your **pdf file** the answers to all of the questions below, including supporting graphs or figures
 - Don't spend too much time on making the document pretty, we expect charts and text to be clear and interpretable by non-data stakeholders, but it does not have to be a polished document/presentation (so screenshots of graphs created in Python or R will do perfectly fine)
 - Please follow the time indicators shown in the task, we expect you to spend at least 50% of your time in section 2B and demonstrate analytical skills especially here
 - You will not be presenting the solution, so make sure we can understand your reasoning from the submission itself
- We do not have additional data points available for this task. Try to complete the questions from the data provided, and if you need to make assumptions, explain your reasoning
- We expect your solutions to be insightful, justified, clearly communicated and understandable by decision makers without a background in statistics.

You can find a **detailed description** of the available data on the last page of this assignment. We wish you the best of luck!

Expectation

We are looking for an entry-level person with a high level of data literacy. You are expected to demonstrate your technical skills and your ability to apply statistical methods in a real-world context, show proficiency in Python, R, or SQL for data extraction, exhibit basic coding skills, and display problem-solving skills and the capability to deliver data-driven insights that align with business objectives.

Use of AI tools

We expect candidates to use AI tools as support, not as a complete task solver:

- The overall approach, assumptions, metric choices, and conclusions should be yours, not generated end-to-end by AI.
- You may use AI for things like:
 - Clarifying concepts or statistical methods.
 - Getting help with language, structure, or small code issues (syntax, debugging).
- Please do not:
 - Ask AI to fully solve the assignment and submit that output as your own.
 - Copy-paste complete AI-generated analysis or code without understanding it.
- At the end of your PDF, include a short note describing how you used AI, if at all (e.g. "Used AI to rephrase some explanations and to fix a pandas groupby error").

We know there is no single 'right' approach to this task. What matters most is the quality of your thinking, your ability to work with data, and how well you communicate your insights with AI playing a supporting role, not the main one.

Introduction

Vinted is a second hand marketplace where *Sellers* upload items and try to sell them to *Buyers*. It is normal that supply is larger than demand, and not all items are being sold. However, our aim is to make sellers happy. In order to provide sellers with as many opportunities to sell as possible, they can buy specific features to promote their items. One of the available promotion features is the "push up" feature that boosts item visibility.

"Push ups" increase the attention given to a specific item on the platform by increasing the visibility for potential buyers (i.e. more views on the homepage and category page) for 3 days after having purchased the "push up" feature. As a result, the item that is up for sale will get significantly more views (on average). Currently, it costs €2 to buy the

“push up” feature. Vinted is considering whether this is the best monetization strategy for the value added services.

Your task is to come up with a data-driven approach to evaluate the existing “push up” monetization strategy and give suggestions on how we could improve the monetization of the platform and its long-term effects.

Section I: Evaluate data quality (Estimate 1 hours)

Before proceeding with the analysis, we should first analyse the dataset's quality:

- A. Perform an exploratory analysis of the dataset provided, and document your observations. How would you describe the data? What are interesting features of the dataset?
- B. How would you perform testing on this dataset to ensure its quality?
- C. What is the granularity of the dataset?
- D. Suggest descriptions for the columns in the dataset. Comment briefly on what you have decided to include or omit in your descriptions.

Section II: Evaluate current situation (Estimate 2 hours)

Product teams at Vinted often introduce new features to the platform. An important step in deciding whether to continue a certain feature is to evaluate the relevant metric(s).

- A. Based on the provided data, what metric can you define to check if users are interested in this feature and why?
- B. Calculate the metric you have suggested in (A). Which categories perform the best according to this metric?
- C. Define another metric that shows a strong correlation with the metric you have found in (A). Discuss the relationship between both metrics using a graph.
- D. In the GIRLS_CLOTHING / FOR_BABIES category, the price of the push up feature is more than half of the average listing price. However, there are still sellers willing to pay for the feature. Why do you think this is the case?

Section III: Analyse alternative strategies (1 hour)

Besides the “push up” feature in its current form, one could think of alternative ways to monetize this feature.

- A. Your colleague suggests setting a fixed price for the “push up” feature different from €2. What implications will a price increase / decrease have on the business, e.g. visibility of listing on the platform?
- B. How would you design a test to validate if a change was successful? What statistical test would you use and why?

Dataset

The dataset is available in the attached file.

Column	Type
category_2	string
category_3	string
number_of_listings	integer
avg_listing_price_eur	float
revenue_from_push_ups	float